Spliceosomal protein and its use

Description

The present invention relates to a spliceosomal protein which is associated with the U11/U12 snRNP complex of the AT-AC spliceosome and is specific for said spliceosome. The invention further relates to the use of said protein and the DNA sequences coding therefor for diagnosing auto-immune diseases and disorders based on defects in the splicing apparatus.

It has been shown in patients suffering from Grave's disease, that incorrect splicing leads to a crucial enzyme (thyroperoxidase) being produced in an inactive form (Zanelli, E. (1990) Biochem. Biophys. Res. Comm., 170, 725). Studies on the spinal muscular atrophy disorder indicate that a defective SMN (survival of motor neurons) gene product leads to considerable disruption of snRNP formation. Inhibition of the splicing apparatus of muscular neurons results in paralysis of the nerve cells and in degradation of muscular tissue (Fischer, U. et al., (1997), Cell, 90: 1023-9; Liu, Q. et al. (1997), Cell, 90: 1013-21; Lefebvre, S. et al., (1997) Nat. Genet., 16, 265). Particular alternative splice variants of the membrane-bound molecule CD44, inter alia, seem to play a crucial part in metastasizing of cancer cells. The CD44 gene comprises a plurality of exons of which 10 adjacent exons are spliced in different order during the generation of mRNA from pre-mRNA. In carcinoma cells of rats it has been detected that metastasizing variants carry exons 4 to 7 or 6 to 7. It has been possible to suppress metastasizing effectively with the aid of antibodies against the part of the protein encoded by exon 6 (Sherman, L., et al., (1996) Curr. Top. Microbiol. Immunol. 213: 249-269).

Incorrect splicing can result in highly distinctive phenotypes of the affected organism. Thus it is known that a point mutation in a β-globin intron can lead to β^{+}- thalassemia. The point mutation results in a wrong splicing location which leads to an altered reading frame and to premature termination of the peptide chain (Weatherall, D.J. & Clegg, J.B. (1982) Cell, 29, 7; Fukumaki, Y. et al. (1982) Cell, 28, 585). In Arabidopsis thaliana mutants, for example, a point mutation in the 5' splice site of the

phytochrome B gene leads to incorrect expression of the gene. Said alteration makes it impossible to remove an intron which contains a stop codon in its sequence. Development of the plant is disrupted, since the gene is involved in phytomorphogenesis (Bradley, J.M. et al. (1995) Plant Mol. Biol. 27, 1133).

Numerous eukaryotic protein- or else RNA-encoding genes have a mosaic-like composition, with individual sequences (exons) which code for part regions of the gene product and intervening sequences (introns) which do not code for the gene product alternating. The primary transcripts of such mosaic genes therefore contain in their RNA chain sequences of both the coding exons and the noncoding introns and have to be processed first for correct gene expression.

One of the essential processing steps is splicing which takes place in the nucleus. In the case of expressing protein-encoding genes, this involves excising intron RNA sequences from relatively long-chain primary transcripts, the “pre-mRNA”, with the formation of mature mRNA and linking exon RNA sequences to one another. The splicing process is catalyzed by a “spliceosome”, a large ribonucleoprotein complex which is assembled in stages from a plurality of small nuclear ribonucleoprotein particles (small nuclear RNPs, snRNPs) which in turn comprise uridine-rich RNAs (small nuclear RNAs, snRNAs) and proteins binding specifically thereto and from proteins, the “non-snRNP splicing factors”, which are not tightly bound to said snRNPs. Splicing generally takes place according to a two-stage mechanism, and at each stage a transesterification step is involved. In the first step, after spliceosome binding to the 5’ splice site and the “branching site” in the intron, a free 5’ exon and a “lariat intron structure” are generated, the intron still being connected to the 3’ exon. In the second step, the two exons are then ligated and the intron is released.

The majority of pre-mRNA introns in metazoa possess terminal, invariable GU dinucleotides and AG dinucleotides. These introns are excised by the “U2-dependent spliceosome” (or major spliceosome) which recognizes said dinucleotides. The spliceosome contains the U1, U2, U5 and U4/U6 snRNPs which are composed of one (U1, U2, U5) or two (U4/U6) RNAs and numerous proteins, for example proteins of the Sm class. The spliceosome recognizes the splice and branching sites of “U2-dependent

introns" by means of interactions in which both RNA and proteins are involved. Thus, various polypeptides, such as the 70K protein and the C protein of U1 snRNPs and proteins of the Ser- and Arg-rich SR-protein family, facilitate duplex formation between U1 snRNA and 5' splice site. Similarly, base pairing between U2 snRNA and the branching site requires numerous U2-snRNP proteins, in particular subunits of the heteromeric splicing factors SF3a and SF3b [R. Reed, Curr. Opin. Genet. Dev. 6, 215 (1996); C.L. Will and R. Lührmann, Curr. Opin. Cell Biol. 9, 320 (1997); A. Krämer, Annu. Rev. Biochem. 65:367 (1996)].

More recently, it has been possible to identify an alternative spliceosome, the "AT-AC spliceosome", which is composed of the U11, U12, U5 and U4atac/U6atac snRNPs and splices an uncommon class of pre-mRNA introns which have AU dinucleotides and AC dinucleotides or GT dinucleotides and AG dinucleotides on their termini [C.B. Burge et al. In: The RNA World II, R.F. Gesteland and J.F. Atkins, eds., Cold Spring Harbour Press, Cold Spring Harbour, N.Y., 1999, p. 525]. These introns are therefore called U12-dependent and, compared with U2-dependent introns, are found only infrequently and contain at the 5' splice site and the branching site highly conserved sequence elements which differ from the weakly conserved sequences of the U2-dependent pre-mRNA introns [C.B. Burge et al. In: The RNA World II, R.F. Gesteland and J.F. Atkins, eds., Cold Spring Harbour Press, Cold Spring Harbour, N.Y., 1999, p. 525; S.L. Hall and R.A. Padgett, J. Mol. Biol. 239:357 (1994); P.A. Sharp and C.B. Burge, Cell 91:875 (1997)]. During assembly of the U12-dependent spliceosome, U11 snRNP binds to the 5' splice site and U12 snRNP binds to the branching site via base pairing [S.L. Hall and R.A. Padgett, Science 271:1716 (1996); W.-Y. Tarn and J.A. Steitz, Cell 84:801 (1996); W.-Y. Tarn and J.A. Steitz, Proc. Natl. Acad. Sci. USA 94:6030 (1997); I. Kolossova and R.A. Padgett, RNA 3:227 (1997)]. The U5 snRNPs and U4atac/U6atac-snRNPs then associate to form the mature spliceosome [W.-Y. Tarn and J.A. Steitz, Science 273:1824 (1996)]. U11 snRNPs and U12 snRNPs are present in nuclear extracts as highly stable 18S U11/U12 complexes [K.M. Wassarman and J.A. Steitz, Mol. Cell Biol. 8, 1276 (1992)], and more recent in vitro binding studies lead to the assumption that U11 and U12 interact as preformed complex with the pre-mRNA [M. Frilander and J.A. Seitz, Genes & Dev. 13:851 (1999)]. These observations, together with the fact that U12 type introns do not have the

essential pyrimidine tract of the 3' splice site of main class introns, leads to the assumption that there are differences between the mechanisms of splice site recognition in the various spliceosomes.

Identification and characterization of proteins characteristic for the U12-dependent spliceosome could therefore not only provide information about the detailed mechanism of the splicing processes in said spliceosome but also make possible at the same time diagnosis and therapy of disorders which can be related to disruptions in the splicing mechanism of said spliceosome. Providing specific proteins makes it further possible to find and develop potential splicing modulators which may also be employed advantageously in the treatment of disorders caused by disruptions of the splicing process.

It was therefore an object of the present invention to provide a protein characteristic for the U12-dependent spliceosome.

This object was achieved by a spliceosomal protein which is associated with the 18S U11/U12 snRNP complex of the U12-dependent spliceosome and which is specific for said spliceosome.

The present invention therefore relates to a spliceosomal protein having the function of the 35kD protein associated with the U11/U12 snRNP complex of the U12-dependent spliceosome, which spliceosomal protein is encoded by a nucleic acid (nucleic acid of the invention hereinbelow) which is selected from the group comprising

a) DNA sequences which encode a protein having the amino acid sequence according to SEQ ID No. 18;

b) DNA sequences which hybridize with the sequences complementary to the sequences under a) and are capable of encoding a protein having the function of the 35kD protein of the U11/12 snRNP complex; and

c) DNA sequences whose genetic code is degenerated with respect to the sequences mentioned under a) or b);

A "spliceosomal protein having the function of the 35kD protein associated with the U11/U12 snRNP complex of the U12-dependent spliceosome (spliceosomal protein of the invention hereinbelow)" herein means any polypeptide which essentially possesses the properties of the naturally occurring, preferably human, 35kD protein of the U11/U12 snRNP complex, i.e. ensures correct functioning of said spliceosome in the spliceosomal complex.

The term hybridization here means a hybridization under customary hybridization conditions, in particular under stringent hybridization conditions, as are known to the skilled worker [Sambrook et al., Molecular Cloning, A Laboratory Manual, 2nd edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989)].

The spliceosomal protein preferably has the amino acid sequence according to SEQ ID No. 18. However, the protein may also have one or more amino acid deletions, amino acid exchanges or amino acid additions or amino acid insertions, so long as the adverse effect caused thereby upon the protein function is negligible. Likewise it is possible, for example, for the spliceosomal protein to contain foreign protein sequences (e.g. as fusion protein).

The preferred inventive spliceosomal protein of 246 amino acids in length has an apparent molecular weight, determined by SDS-Page, of approximately 35kD, a calculated molecular weight of 29kD and an isoelectric point of 9.88.

Another subject are the DNA sequences as claimed in claim 1, which code for said protein and sequences complementary to said sequences and also fragments thereof.

Said DNA sequences may be linked to other DNA sequences, in particular sequences which make it possible to express the protein in a desired host organism. Sequences of this kind are known in the prior art. They may be, for example, regulatory sequences such as promoter sequences, Shine-Dalgarno sequences, transcription termination signals, polyadenylation signals or enhancer elements. In this way it is possible to produce the desired protein in large amounts at low cost.

The invention therefore also relates to recombinant DNA molecules which contain the DNA sequences of the invention.

The recombinant DNA molecules may either be introduced directly into the desired host organism or first be incorporated into vectors which are then used to transform the host organisms in a manner known per se. The invention likewise relates to vectors of this kind. Vectors which may be used are the vectors customary in the prior art, for example plasmids, bacteriophages or viruses. Preferred vectors are expression vectors.

The invention likewise relates to host organisms which contain the recombinant DNA molecules or vectors of the invention.

Examples of suitable host organisms are prokaryotic or eukaryotic micro-organisms, for example bacteria such as Escherichia coli, yeasts or tissue cells.

The DNA sequences of the invention or fragments thereof may be used for finding homologous DNA sequences in various organisms or tissue types which possess a similar or identical function to the spliceosomal protein of the invention.

The protein of the invention and the DNA sequences coding for said protein furthermore may be [lacuna] advantageously as diagnostic agents, for example for autoimmune diseases and of disorders which can be related to disruptions in the splicing mechanism.

Thus it is known that spliceosomal components may act as autoantigens. patients suffering from the autoimmune disease systemic lupus erythematosus, for example, frequently produce antibodies which can be used to precipitate the majority of snRNPs. By way of a simple immunoassay, the protein of the invention now provides a possibility for rapid diagnosis of autoimmune diseases which are based on antibody formation against proteins which are specific for the U12-dependent spliceosome.

Disorders which can be related to disruptions in the splicing mechanism of the U12-dependent spliceosome, which disruptions are based on a defect in the 35kD protein, may be diagnosed, for example, with the aid of

complementation assays in in vitro splicing systems known in the prior art. Here, first a pre-mRNA having a U12-dependent intron which contains all structural elements necessary for recognition of the pre-mRNA by the spliceosome and for the splicing process is prepared, for example by in vitro transcription. The RNA is, for example, radiolabeled so that it is later, after fractionation on a denaturing urea polyacrylamide gel, possible to assess, on the basis of the characteristic band patterns, whether a splicing reaction has occurred. Subsequently, samples from patient tissue are assayed with and without addition of the 35kD protein of the invention. Complementation then confirms the defect in said protein.

Another usable in vitro splicing system is described in PCT/EP 00/01595.

The DNA sequences of the invention may be employed by way of customary hybridization assays for the diagnosis of defects in the gene for the described 35kD protein, in particular in prenatal diagnostics.

The protein of the invention may further be used as a therapeutic agent for disorders based on splicing defects.

It is further possible to use the spliceosomal protein of the invention advantageously for finding or developing splicing modulators, for example splicing inhibitors, which are then suitable for the treatment of further disorders. Thus, most recently a U12-dependent intron has been identified in the mutant genes which are responsible for the autosomally recessive Hermansky-Pudlak-Syndrome (HPS). It can be expected that in the future such introns may also be found in further genes to which a part in genetically caused disorders can be attributed. A specific inhibition of the splicing of introns present in such genes and thus the expression of the harmful mutated genes is therefore a possible way of treating said disorders. Due to the rarity of U12-dependent introns, it is also possible to use such splicing inhibitors therapeutically more effectively for U12-dependent spliceosomes than for U2-dependent spliceosomes.

The sought after splicing modulators may be found by using the known in vitro assay systems already mentioned above for studying splicing mechanisms. Examples of splicing modulators or splicing inhibitors which can be analyzed in this way are monoclonal antibodies. Thus, influencing

splicing processes in the cell, for example generation of mature mRNA, with the aid of antisera or monoclonal antibodies has already been described [R.A. Padgett et al., Cell 35:10 (1983); R. Gattoni et al., Nucleic Acid Res. 24:2535 (1996)].

The identification of the regulatory processes in splicing by the minor spliceosome is important for virology, too. Recently (Hibbert et al., 1999), a sequence of the U1 5' splice site in a region of the negative regulatory splicing element (NRS) in Rous sarcoma virus has been identified. Binding of U1 snRNPs to the NRS inhibits the splicing process, i.e. increases NRS activity.

The U1 binding site overlaps with a U11 binding site which is identical to a sequence on the NRS in the 5' splice site of the U12-dependent spliceosome. U11 snRNP binding to the NRS promotes splicing, i.e. reduces NRS activity. The competition between U1 snRNP and U11 snRNP contributes to establishing equilibrium between spliced and unspliced viral RNAs for the purpose of optimal viral replication.

As studies by Hibbert et al. (1999) show, specific mutations in the U1 and/or U11 binding domain would be able to contribute to influencing and possibly even controlling viral life cycles. A strategy for inhibiting U11/NRS interaction could be developed by means of information relating to the pre-spliceosomal characteristic, which was obtained in this study.

The invention therefore also relates to a pharmaceutical comprising a nucleic acid of the invention and/or a spliceosomal protein of the invention and, where appropriate, pharmaceutically acceptable additives and/or excipients.

Furthermore, the invention relates to a process for producing a pharmaceutical for the treatment of cancer, autoimmune diseases, in particular Grave's disease, spinal muscular atrophy, β'-thalassemia, cancers related to the c-erb oncogene, hepatitis C infection, herpes simplex virus infection, systemic lupus erythematosus, Hermansky-Pudlak syndrome, with a nucleic acid of the invention or a spliceosomal protein of the invention being formulated.

The invention likewise relates to a diagnostic agent comprising a nucleic acid of the invention and/or a spliceosomal protein of the invention and, where appropriate, pharmaceutically acceptable additives and/or excipients.

Furthermore, the invention relates to a process for producing a diagnostic agent for diagnosis of Grave's disease, spinal muscular atrophy, β'-thalassemia, cancers related to the c-erb oncogene, hepatitis C infection, herpes simplex virus infection, systemic lupus erythematosus, Hermansky-Pudlak syndrome, with a pharmaceutically acceptable carrier being added to a nucleic acid of the invention and/or a spliceosomal protein of the invention.

Description of the figures and most important sequences:

Fig. 1 shows the snRNA composition of purified snRNPs.

Fig. 2 shows the protein composition of U11/U12 snRNPs selected using oligonucleotides.

SEQ ID No. 17 shows the genomic DNA sequence for the U12-associated 35kD protein including noncoding sequences. The coding sequence is indicated by the derived amino acid sequence below and corresponds to SEQ ID No. 18.

SEQ ID No. 18 shows the amino acid sequence of the U12-associated 35kD protein, comprising 246 amino acids.

The isolation of the U11/U12-associated 35kD protein of the invention is illustrated below by way of example.

The following examples serve to illustrate the invention in more detail without restricting it to said examples.

Examples

Preparation of HeLa nuclear extracts

Nuclear extracts from HeLa cells were prepared by culturing cell cultures with HeLa cells. The cells were then sedimented from the culture medium by centrifugation (1 000 × g, 10 min.) and washed with phosphate buffer. This was followed by taking up the cell sediment in five volumes of buffer A (10 mM HEPES, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT, pH 7.9, 4°C) and incubating for 10 minutes. The cells were again sedimented and taken up in two volumes of buffer A. This suspension was disrupted using a Dounce homogenizer (plunger B) (moving the plunger up and down 10 times). The nuclei were sedimented by centrifugation. The nuclei were then again taken up in buffer A and centrifuged at 25 000 × g for 20 minutes. The sediment was taken up in 3 ml of buffer B (20 mM HEPES, 25% (v/v) glycerol, 0.42 M NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM phenylmethylsulfonyl fluoride (PMSF), 0.5 mM DTT, pH 7.9) and again disrupted using the Dounce homogenizer. The resulting suspension was incubated on a magnetic stirrer for 30 minutes and then centrifuged at 25 000 × g for 30 minutes. This was again followed by centrifugation at 25 000 × g (30 min.). The supernatant was dialyzed against 50 volumes of buffer C (20 mM HEPES, 20% (v/v) glycerol, 0.1 M KCl, 0.2 mM EDTA, 0.5 mM PMSF, 0.5 mM DTT, pH 7.9). The dialysate was centrifuged (25 000 × g, 20 min.). The resulting superantant may be stored as nuclear extract in liquid nitrogen (Dignam, J.D. et al. (1983) Nucleic Acid Res., 11, 1475).

Isolation and analysis of snRNPs

Spliceosomal snRNPs which had been provided beforehand with a trimethylguanosine (m3G)-cap were purified from nuclear extracts of HeLa cells via immunoaffinity chromatography using anti-m3G antibodies and fractionated by sedimentation through a 10-30% strength glycerol gradient [B. Laggerbauer, J. Lauber and R. Lührmann, Nucleic Acids Res. 24:868 (1996)]. Fractions containing the 18 S U11/U12 snRNP complexes were pooled and the KCl concentration was adjusted to 250 mM. snRNPs from 2.4 ml of the pooled 18S gradient fractions were incubated for 16 h at 4°C with 12 μg of oligonucleotide which is complementary either to nucleotides 2-18 of human U11 snRNA,

5'ACGACAGAAGCCCUUUUdT*dt*dT*dT*-3' (U11 oligo). or to nucleotides 11-28 of human U12 snRNA, 5'-AUUUUCCUUACUCAUAAG-dT*dT*dT*dT*-3' (U12 oligo), with * being a biotinylated 2'-deoxythymidine and A, U, G and C being 2'-O-methylribonucleotides. The oligonucleotide-bound snRNPs were precipitated using streptavidin agarose in a manner known per se [A. I. Lamond and B. S. Sproat, in RNA Processing: A Practical Approach, D. Rickwood and B. D. Hames, eds. (Oxford University Press, Oxford, 1996) p. 103]. RNA from 1/5 of agarose-precipitated snRNPs was eluted by incubation for 30 minutes at 95°C in 100 µl of DH buffer (15 mM NaCl, 1.5 mM Na-citrate, 0.1% SDS), fractionated on 10% strength polyacrylamide/7M urea gels and made visible by silver staining. The identity of the selected RNAs was confirmed as U11 and U12 by Northern blotting. The protein was eluted from the remaining beads by incubation at 95°C in 200 µl of S buffer (60 mM Tris, pH 6.8, 1 mM EDTA, 17.5% glycerol, 2% SDS, 0.2 M DTE) for 5 minutes and precipitated with 5 volume units of acetone. The proteins were fractionated via SDS-PAGE on gels with 10% (top half) and 13% (bottom half) polyacrylamide and made visible by Coomasie staining. For comparison, RNA and protein from 50 µl of the starting material (pooled 18S gradient fractions) were also analyzed.

The results for oligo U11 are depicted in Fig. 1, lane 2, and those for oligo U12 in Fig. 1, lane 3. Lane 1 shows the snRNAs of the starting material (input); lane 4 shows the control for precipitation in the absence of oligonucleotides (mock). Coselection of U12 with an oligonucleotide directed against U11 and vice versa showed that mainly U11/U12 snRNP complexes and not U11 monoparticles or U12 monoparticles had been selected. Accordingly, Fig. 2, lanes 2 and 3, shows identical protein patterns of U11/U12 snRNPs, independent of which oligonucleotide had been used to select said snRNPs. The molecular weight of the proteins (in kD) is indicated on the right. Lane 1 again shows the proteins of the starting material (input) and their identity (left-hand side). Lane 4: control.

Identification of proteins associated with the U11/U12 snRNP complex

Of the 20 different proteins found in the U11/U12 complex, 8 migrated with the Sm snRNP nuclear proteins B', B, D3, D2, D1, E, F and G, which are present in the major spliceosome snRNPS (Fig. 2, lanes 1-3). Antibodies which reacted specifically with B'/B, D3, D2, F or G recognized even the

proteins having identical molecular weights on immunoblots of the U11/U12 complex. U11/U12 therefore contains the same 8 Sm proteins which are found in the major spliceosome too.

Of the 12 remaining U11/U12 proteins, the 160kD, 150kD, 130kD and 49kD proteins of the U11/U12 complex migrated together with four of the proteins specific for the 17S U2 complex, which make up the essential splicing factor SF3b, namely U2-160, U2-150, U-120 and U2-53. Antibodies directed against U2-160, U2-150 and U2-120 moreover reacted strongly with the 160kD, 150kD and 130kD proteins of the U11/U12 complex. Peptide sequences which had been obtained by microsequencing of the 160kD, 150kD, 130kD and 49kD proteins of the U11/U12 complex (Table I) proved to be essentially identical with the known SF3b sequences. The four proteins mentioned are thus very likely identical to the proteins known from SF3b, and deviations in the apparent molecular weights of some of the comigrating proteins can be attributed to differences in the electrophoresis conditions which were applied in the original identification of U2 snRNP proteins.

Table 1

U11/U12 Protein	**Peptides**
160kD	KMNARTYMDVMREQHLTK KLTATPTPLGGMTGF KAIVNVIGMH
150kD	KRIFEAFK KLRRMNRFTVAE KRTGIQEMREALQEK KLTIHGDLYYEG
130kD	KLGAVFNQVAFPLQYT KLLRVYDLGK KNVSEELDRTPPEVSK KLENIAQRYAF
49kD	KVSEPLLXELFLQ KDRVTGQHQGYGFVEFLSEE

X here means any amino acid.

Characterization of the 35kD protein associated with the U11/U12 snRNP complex

Of the remaining proteins, the 35kD protein associated with the U11/U12 snRNP complex was characterized. To this end, the peptides KEYDPLK and KRWQEREPTRVWPDND were obtained by micro-sequencing the 35kD protein which had been fractionated on a gel and digested with trypsin. These peptides were used to screen the EST database of the National Center for Biotechnology Information for cDNAs by means of the TBLASTN program. Both peptides were found in an ORF of a cDNA from human macrophage cells (Genbank Accession No. U44798) which codes for an unknown protein. A second cDNA from human muscle cells in pBluescript SK, containing an identical ORF (Genbank Accession No. AA211268) was transformed into E. coli HB101 and then fully sequenced using an ABI PRISM sequence analyzer. The DNA sequence coding for the 35kD protein is depicted in SEQ ID No. 17 together with the amino acid sequence derived therefrom. SEQ ID No. 17 inherently shows the complete cDNA sequence including noncoding 5' and 3' sequences.

Protein was prepared by in vitro translation of the cDNA (TNT (coupled transcription/Translation) kit from Promega). The protein migrated on an SDS polyacrylamide gel together with the purified 35kD protein, proving that the DNA encodes a complete protein.

The U11/U12 35kD protein has an RNA recognition motif (RRM; amino acids 51-129). This region and the adjacent glycine-rich region are very similar to a region of the U1 70K protein. Moreover, antisera against the 35kD protein efficiently precipitated U11 from a mixture of snRNPs fractionated through a gradient and containing U11 monoparticles. In analogy to U1 70K, the U11 35kD protein therefore might facilitate 5' splice site recognition. Furthermore, the protein could be involved in exon linking, interacting with components of the major spliceosome.